

H.264 표준을 위한 효율적인 인트라 예측 모드 결정 방법

준회원 박지윤*, 정회원 이창우*

An Efficient Intra Prediction Mode Decision Method for H.264 Standard

Ji-yoon Park* Associate Member, Chang-woo Lee* Regular Member

요약

가장 최근에 표준화된 H.264 동영상 압축 부호화 기법에서는 많은 새로운 기술들을 적용하여 높은 부호화 성능을 보인다. 특히 최적의 인트라 예측 모드를 결정하기 위해 모든 예측 모드들에 대해 비트율 왜곡 최적(rate distortion optimization) 기법을 적용하기 때문에 많은 계산량을 필요로 한다. 본 논문에서는 분산을 이용하여 인트라 모드를 결정하고 영상의 edge들을 찾아내는 연산을 통해 블록의 방향성을 찾고 그 방향성을 이용하여 인트라 예측 모드 결정에 필요한 계산량을 감소시키는 효율적인 기법을 제안한다.

Key Words : Video coding, H.264, intra prediction, Mode decision

ABSTRACT

The H.264/AVC video coding standard shows superior coding efficiency by adopting many new techniques. However, the encoding complexity increases greatly to achieve higher coding efficiency. Especially, the rate distortion optimization technique, which is used to decide the intra-prediction mode, increases the encoding complexity. In this paper, we propose an efficient intra-prediction mode decision method. By using the variance of pixel values and the edge direction, the computational complexity of the intra-prediction mode decision is greatly reduced.

I. 서론

H.264 표준 부호화 방식은 기존의 비디오 부호화 표준들인 MPEG-2, MPEG-4 와 H.263 등에는 없던 새로운 부호화 기법들을 적용하여 부호화 효율을 높인다^[1]. 예를 들면, 인트라 코딩에서 정교한 공간적 예측, 가변 블록 크기 움직임 보상, 4x4 정수변환, 복수 참조 픽처, 그리고 문맥 기반 적응적 이진 산술 부호화 방식(CABAC) 등의 새로운 부호화 기법들이 적용되었다^[2].

H.264 부호기는 최적의 인트라 모드 결정 및 인트라 예측 모드 결정을 위해 모든 모드에 대해서 비트율 왜곡

비용 값(rate-distortion cost)을 계산한다. 즉, 인트라 4x4 모드에서는 9 가지의 모드, 그리고 인트라 16x16 모드에서는 4 가지의 모드에 대해서 비트율 왜곡 최적 기법이 적용되어 그 비용 값을 계산한다. 따라서, 최적의 모드를 선택하기 위해 요구되는 계산량 및 수행 시간이 상당히 증가하게 되고, 인트라 예측 모드를 결정하기 위한 효율적인 방법에 대한 많은 연구가 진행되고 있다^[4-10]. 이를 위해서 Pan 등은 소벨 연산(Sobel operator)를 이용한 edge direction histogram 방법을 토대로 선택 가능한 예측 모드의 수를 줄임으로서 인트라 모드 선택 과정에서 복잡도를 감소시키는 방법을 제안하였다^[7]. Wang 등은 DES(dominant edge strength)를 이용하여

※ 본 연구는 2008년도 가톨릭 대학교 교비 연구비 지원으로 이루어졌습니다.

* 가톨릭대학교 정보통신전자공학부 신호처리 연구실 (wmdnls79@catholic.ac.kr, changwoo@catholic.ac.kr)

논문번호 : KICS2008-02-109, 접수일자 : 2008년 2월 28일, 최종논문접수일자 : 2008년 9월 8일

인트라 4x4 모드에서는 9 가지 모드 중에서 4 가지 모드를 택하였고 인트라 16x16 모드에서는 4 가지 모드 중 2 가지 모드를 먼저 선택하여 인트라 모드 선택 과정에서 복잡도를 감소시키는 방법을 제안하였다⁸⁾.

본 논문에서는 인트라 모드를 선택할 때 복잡성을 감소시키기 위해 공간 영역에 매크로 블록의 분산 (variance) 값을 계산하여 인트라 모드를 결정하고 16x16 혹은 4x4 인트라 모드에서 로버트 연산(Roberts operator)를 이용한 gradient 값을 이용하여 선택 가능한 예측 모드의 수를 감소시키는 효율적인 인트라 예측 모드 결정 방법을 제안한다.

본 논문의 구성은 다음과 같다. 먼저 II장에서는 인트라 모드 결정을 위해서 참조 소프트웨어(JM 11.0)에서 사용하는 방법과 기존의 방법들에 대하여 기술하고, III장에서는 본 논문에서 제안하는 고속 인트라 모드의 선택 방법에 대해 구체적으로 설명한다. 그리고 IV장에서 기존의 방법들과 제안하는 방법과의 비교 분석을 위해 다양한 크기의 영상들을 가지고 실험한 결과 값을 제시한다. 마지막으로 V장에서 결론을 맺는다.

II. 참조 소프트웨어를 통한 인트라 모드 결정 방법과 기존의 알고리즘

2.1 참조 소프트웨어를 통한 인트라 모드 결정 방법

H.264 부호화 기법에서 일반적으로 인트라 4x4 모드는 복잡한 매크로 블록에 대해서 사용하는 것에 적합한 반면, 인트라 16x16 모드는 전체적으로 영상의 단순로운 매크로블록일 경우에 사용하는 것이 적합하다. 인트라 4x4 모드를 사용할 때, 그림 1(a)와 같이 각각의 예측되어야 할 4x4 휘도 블록은 공간적으로 이웃하는 화소 값으로부터 구해지며 그림 1(b)에 9 가지 예측모드에 대한 방향성을 도시하였다. 또한, 인트라 16x16 모드에는 그림 2와 같이 4 가지 예측모드 중 DC 모드를 제외한 3 가지 예측모드 방향이 존재한다. 그림 3에 나타난 것처럼 참조 소프트웨어에서 비트율 왜곡 최적화 기법을 사용하는 경우 각 예측모드에 대해서 인트라 16x16 모드와 인트라 4x4 모드의 비트율 왜곡 비용 값을 비교하여 최소의 비용 값을 갖는 모드를 최적의 모드로 선택한다. 비트율 왜곡 비용함수는 식 (1)과 같이 계산될 수 있다.

$$RDcost = Distortion + \lambda \times Rate \quad (1)$$

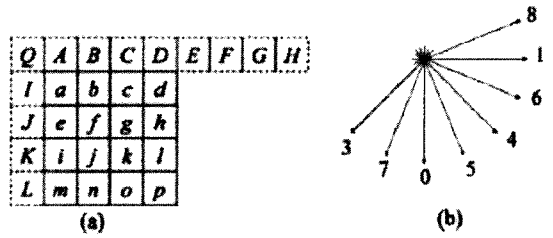


그림 1. (a) 인트라 4x4 블록 및 예측방향 4x4 휘도블록(a-p) 및 예측에 사용되는 샘플(A-Q), (b) 인트라 4x4모드의 9가지 예측모드 중 DC모드를 제외한 8가지모드의 예측방향

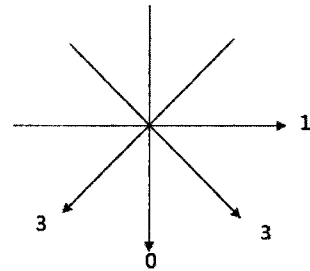


그림 2. 인트라 16x16모드의 DC모드를 제외한 3가지모드의 예측방향

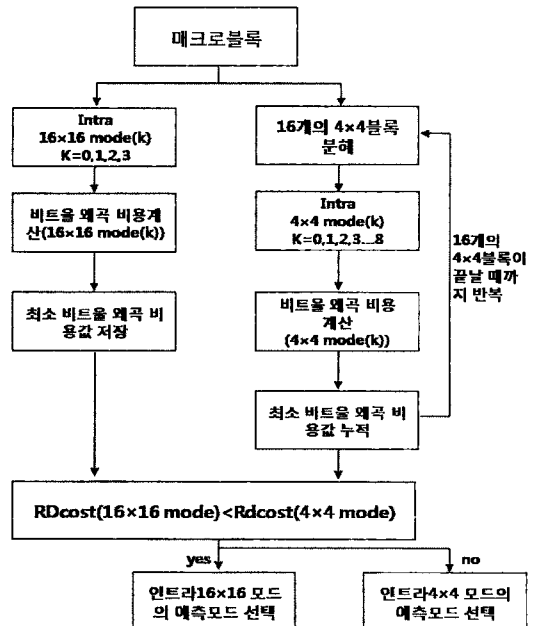


그림 3. edge의 인트라 예측과정

식 (1) 에서 인트라 4x4 모드는 각 4x4 블록 마다 9 가지 예측모드로부터 비트율 왜곡 비용 값을 계산 하며 그 중에서 최소 값을 각 블록마다 선택하게 된다. 또한 인트라 16x16 모드는 매크로블록 전체에서 4 가지 예측모드에 대한 비트율 왜곡 비용 값들 중 최

소 값을 가지는 모드로 선택할 수 있다. 따라서 인트라 4×4 모드에 의해 얻어진 16개의 4×4 블록의 비트율 왜곡 비용 값의 합과 인트라 16×16 모드에 의해 얻어진 비트율 왜곡 비용 값을 비교하여 작은 값을 갖는 모드를 최적의 인트라 모드로 선택한다.

따라서 하나의 매크로블록에 대한 최적의 인트라 모드를 선택하기 위해서는 모든 예측 모드에 대해서 비트율 왜곡 비용 값을 계산하면 최대 592(=4×(1×4+16×9))회의 비트율 왜곡 비용 계산이 필요함을 알 수 있다. 이와 같이 모든 모드에 대해 비트율 왜곡 최적 기법을 적용하여 최적의 모드를 선택하는 것은 많은 연산을 필요로 하므로 계산량을 줄이기 위한 여러 가지 기법들이 제안되고 있다³⁻¹⁰⁾.

2.2 인트라 예측 모드를 효율적으로 결정하기 위한 기존의 방법

2.1절에서 설명한 것과 같이 인트라 예측 모드 결정을 위해서 요구되는 많은 계산량을 줄이기 위해 여러 가지 해결 방안들이 제시되고 있다. 그중에 대표적인 방안으로 Pan 등이 제안하는 소벨 연산을 이용한 edge direction histogram 방법이 있다⁷⁾. 이는 간단한 edge 검출 알고리즘과 local gradients를 바탕으로 하고 있는데 이러한 방법을 이용하는 이유는 인트라 4×4 모드의 9 가지 예측모드와 인트라 16×16 모드의 4 가지 예측모드가 방향성에 따라 결정되기 때문이다. Pan 등은 방향성을 검출하기 위해서 다음 식 (2)에 나타난 소벨 연산을 사용하였다. 식 (2)에서 $p_{i,j}$ 은 화소 값을 나타내며, $dx_{i,j}$ 과 $dy_{i,j}$ 은 행 방향으로 화소간의 차이 값과 열 방향으로의 화소간의 차이 값을 각각 나타낸다.

$$dx_{i,j} = p_{i-1,j+1} + 2 \times p_{i,j+1} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i,j-1} - p_{i+1,j-1} \quad (2)$$

$$dy_{i,j} = p_{i+1,j-1} + 2 \times p_{i+1,j} + p_{i+1,j+1} - p_{i-1,j-1} - 2 \times p_{i-1,j} - p_{i-1,j+1}$$

Gradient 값을 이용하면 edge 벡터의 크기 및 방향성을 다음 식 (3)과 (4)를 통해 구할 수 있다.

$$Amp_{i,j} = |dx_{i,j}| + |dy_{i,j}| \quad (3)$$

$$Ang_{i,j} = \frac{180^\circ}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right), |Ang_{i,j}| < 90^\circ \quad (4)$$

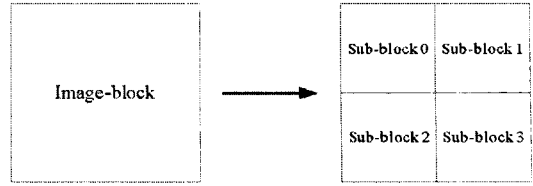


그림 4. 하위블록의 정의

위의 식들을 이용하여 인트라 4×4 모드에서는 4×4 블록 크기에 있는 화소 값들에 대해 소벨 연산을 적용하고 식 (3)과 (4)에서 얻어지는 값들의 DC 모드를 제외한 8 가지 예측 모드의 방향성을 고려하여 경계치를 설정하고 설정된 조건에 맞는 모드에 누적 값을 계산해 나간다. 이 누적 값을 통해 가장 높은 값의 예측 모드와 인접한 예측 모드 두 가지, 그리고 DC 모드를 예측 후보 모드로 선택한다. 인트라 4×4 모드와 같은 방법으로 인트라 16×16 모드에서도 누적 값을 통해 4 가지 예측 모드 중 가장 큰 값을 갖는 하나의 예측 모드와 DC 모드를 예측 후보 모드로 결정하게 된다. 이와 같이 소벨 연산을 이용하여 인트라 예측 모드 결정을 제안한 Pan 등의 연구는 인트라 예측 모드 결정 방법에 대해서 다루었다⁷⁾.

또한 인트라 예측 모드 결정을 위해서 Wang 등은 dominant edge strength 방법을 사용하였다⁴⁾. Wang 등이 제안하는 dominant edge strengths(DES)는 다음과 같이 구한다. 먼저 그림 4와 같이 인트라 예측을 할 블록을 하위블록 4 개로 나누어 각 하위 블록에 속하는 화소 값들의 평균을 구하고 가장 일치하는 방향성을 찾기 위해 5 가지 방향성으로 구분하여 평균값과 5 가지 형태의 필터 계수를 곱함으로써 각각의 DES 값을 구한다. DES 값들 중 가장 큰 값을 갖는 DES 값의 방향성으로 예측 후보 모드를 선택한다. 이와 같이 Wang 등이 제안하는 방법은 전체 모드에 대해 비트율 왜곡 최적 기법을 적용하였던 H.264 표준안 부호기외는 다르게 가능성 있는 모드를 선택하여 모드 수를 한정함으로써 수행 시간을 줄이는 방법이다.

위에 제시한 기존의 방법들의 경우 인트라 예측모드를 선택하고 그에 따라 인트라 모드가 결정되어진다. 이에 비하여 Song 등은 먼저 인트라 모드를 선정하고 선택된 인트라 모드에 대해서 예측 모드를 결정하는 방법을 제안하였다¹⁰⁾. Song 등이 제시한 방법은 다음의 식 (5)에 제시한 SAD(Sum of Absolute Difference)을 이용하여 인트라 모드를 우선순위로 선택한다.

$$T = \sum_{i=1}^{16} \sum_{j=1}^{16} \text{abs}(x_{i,j} - \bar{x}) \quad (5)$$

인트라 4×4 모드를 선택할 것인지 아니면 인트라 16×16 모드를 선택할 것인지에 대한 경계 값의 범위를 설정해 주기 위하여 다음의 식 (6)을 이용한다.

$$T' = \sum_{i=1}^n \frac{r_i}{r_1 + r_2 + r_3 + \dots + r_n} T_i \quad (6)$$

식 (6)에서 n은 부호화하는 전체 프레임 수이며 T_i 는 i번째 해당하는 프레임의 T 값들의 평균이다. 또한 r_i 는 부호화 각 프레임에 대한 전체 프레임의 상관도이다. 식 (6)과 같이 경계치를 설정해 주었을 때 $T_1 = T' - 200$, $T_2 = T' + 200$ 경계치를 이용하여 인트라 모드를 선택한다. 그러나 위 방법은 참조 소프트웨어(JM 11.0)에서 각각의 매크로블록에 대하여 인트라 예측을 수행하는 방법과는 달리 부호화하려는 프레임들의 전체적인 정보들을 먼저 확인하고 연산 수행을 해야만 하는 문제점이 발생하게 된다. 이는 인트라 모드 선택의 경계 값 범위를 설정해 주는데 있어서 많은 계산량이 필요하기 때문에 기존 방

법들에 비하여 높은 효율을 기대할 수가 없다.

III. 공간영역상의 효율적인 인트라 모드 결정

본 논문에서는 매크로 블록의 분산 값을 이용하여 인트라 4×4 모드와 인트라 16×16 모드 중에서 한 모드를 먼저 선택하고 선택된 인트라 모드에서 선택 가능한 예측 모드를 한정하여 일부 모드에 대해서만 비트율 왜곡 최적 기법을 수행하는 효율적인 인트라 예측 모드를 결정하는 방법을 제안한다. 제안하는 방법의 순서도를 그림 5에 도시하였다.

그림 5의 순서도에서 인트라 4×4 모드와 인트라 16×16 모드 중 어느 모드를 선택할 것인지를 결정한다. 식 (7)에 의해 계산되는 매크로 블록의 분산 값을 구하면 매크로 블록의 복잡도를 계산할 수 있다.

$$\sqrt{\text{Var}} = \sqrt{\frac{1}{256} \left(\sum_{i,j=0}^{15} p_{i,j}^2 - 256 \bar{p}^2 \right)} \quad (7)$$

이를 바탕으로 표 1과 표 2는 참조 소프트웨어(JM 1.0)에서 선택된 인트라 모드 16×16 혹은 4×4

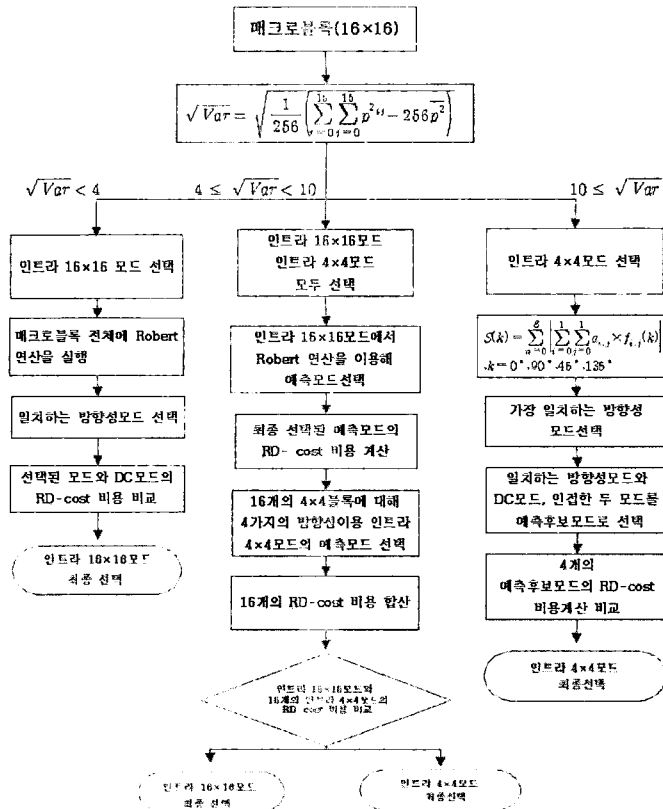


그림 5. 제안하는 알고리즘 순서도

표 1. 인트라 16×16모드를 선택한 매크로 블록에 대한 분산 값 분포 수치

파일명	인트라 16×16선택 매크로블록 수 / 전체 프레임 수	M<4		4≤M<10		10≤M	
		개수	백분율	개수	백분율	개수	백분율
carphone	1727/100	1295	75.0	223	12.9	209	12.1
miss_am	3802/ 75	3619	95.2	130	3.4	53	1.4
foreman	571/100	331	58.0	71	12.4	169	29.6
claire	5179/100	3648	70.5	346	6.7	1182	22.8
stefan	3422/ 45	2969	87.0	188	5.5	265	7.7
football	3017/ 45	1594	53.0	1191	39.5	232	7.7
foreman	3318/ 45	1931	58.2	1077	32.5	310	9.3
news	561/ 45	4279	76.3	760	13.5	572	10.2

표 2. 인트라 4×4모드를 선택한 매크로 블록에 대한 분산값 분포 수치

파일명	인트라 4×4선택 매크로블록 수 / 전체 프레임 수	M<4		4≤M<10		10≤M	
		개수	백분율	개수	백분율	개수	백분율
carphone	8173/100	187	2.3	1039	12.7	6947	85.0
miss_am	3623/75	283	7.8	988	27.3	2352	64.9
foreman	9329/100	42	0.5	669	7.1	8618	92.4
claire	4724/100	43	0.9	665	14.1	4016	85.0
stefan	14398/45	163	1.1	479	3.3	13756	95.6
football	14803/45	727	4.9	6453	43.6	7623	51.5
foreman	14502/45	251	1.7	2897	20.0	11354	78.3
news	12209/45	370	3.0	1744	14.3	10095	82.7

모드에 대한 분산 값 분포 수치를 나타낸다. 45~100 개 프레임의 실험 영상을 대상으로 분산 값의 분포를 계산하여 표시하였으며 표 1과 표 2의 결과는 인트라 모드를 결정할 때 분산 값이 중요한 결정요소가 될 수 있다는 것을 보인다.

참조 소프트웨어에서 분산 값이 4 미만을 가질 때 인트라 16×16 모드를 선택하는 확률이 높다는 것을 표 1을 통해 알 수 있고, 분산 값이 10 이상을 가질 때 인트라 4×4 모드를 선택하는 확률이 높다는 것을 표 2를 통해 확인할 수 있다. 그러나 분산 값 분포를 살펴보면 Foreman QCIF 영상과 Claire QCIF 영상에서 인트라 16×16 모드를 선택할 때 분산 값이 4 미만을 가지는 확률이 다른 영상에 비하여 낮은 편이다. 그러나 Foreman QCIF 영상과 Claire QCIF 영상은 인트라 16×16 모드를 선택하는 비율이 전체에서 5% 내외로 작다. 따라서 인트라 16×16 모드를 선택할 때 분산 값이 10 이상의 값을 갖는 비중이 커더라도 전체 성능에 영향을 미치지 못한다. 또한 분산 값이 4이상에서 10 미만의 사이에 존재할 때는 4×4 모드 혹은 16×16 모드를 결정하기 어렵기 때문에 인트라 모드를 정하지 않고

본 논문에서 제안하는 방법을 사용하여 예측 후보 모드의 수를 줄이는 방법을 사용한다. 따라서 분산 값이 4미만일 때는 인트라 16×16 모드를 선택하고 분산 값이 4이상이고 10미만일 때는 인트라 16×16 모드와 인트라 4×4 모드에 대한 선택된 예측 후보 모드 중에서 결정한다. 그리고 분산 값이 10이상일 때는 인트라 4×4모드를 선택하는 방법을 사용하였다.

인트라 모드를 결정한 후에는 각 인트라 모드에서 예측 모드를 결정하는데 효율적인 방법을 사용해야 한다. 본 논문에서는 인트라 16×16 모드의 경우 로버트 연산(Robert operation)을 사용하여 edge의 방향성을 구하고 이를 통해 가능한 모드 수를 줄이는 방법을 사용하였고, 인트라 4×4 모드의 경우는 네 가지 방향성을 정하고 각 방향성에 따라 다른 가중치를 적용하여 다섯 가지 방향에 대해서만 고려함으로써 각각의 계산량을 줄이는 방법을 제안한다.

인트라 16×16 모드의 4 가지 예측 모드 중 최적의 모드를 선택하기 위한 방법은 다음과 같다. 기존에 사용한 소벨 연산의 경우 각 화소마다 많은 연산량을 필요로 한다. 인트라 16×16 모드의 경우에는 화소의 이웃하는 대각선상의 화소들과의 차이를 이용하여 간단한 gradient 값을 구하는 로버트 연산을 사용하여 계산량을 줄일 수 있다. 식 (8)에 의해 gradient 값을 구하고 edge의 크기와 방향을 식 (9)와 (10)에 의해 구할 수 있다.

$$dx_{i,j} = p_{i,j} - p_{i-1,j+1} \quad (8-a)$$

$$dy_{i,j} = p_{i,j} - p_{i-1,j-1} \quad (8-b)$$

$$Amp_{i,j} = |dx_{i,j}| + |dy_{i,j}| \quad (9)$$

$$Ang_{i,j} = \frac{180^\circ}{\pi} \times \arctan\left(\frac{dy_{i,j}}{dx_{i,j}}\right), |Ang_{i,j}| < 90^\circ \quad (10)$$

예측 모드를 정확히 선택하기 위해서는 edge의 방향성을 나타내는 각도 간의 경계 값을 결정해야 한다. 그림 2와 같이 각 예측 모드가 가지는 각도 간의 간격이 45도라는 것을 고려하여 그 절반인 22.5도씩 경계 값으로 정하였다. 설정된 경계 값들을 이용하여 edge의 방향성에 따라 예측 모드를 선택한다. 즉, 식 (11)에 제시된 행방향의 로버트 연산 값과 열방향의 로버트 연산 값 간의 비율을 통해 영상의 edge 성분을 분석할 수 있고 이를 방향성과 연관시킴으로써 식 (12),(13)과 식 (14)와 같이 화소 간의 차이를 누적하여 예측 모드를 결정할 수 있다.

$$\eta = \frac{dy_{i,j}}{dx_{i,j}} \quad (11)$$

$$histo(0) += Amp(\overrightarrow{D_{i,j}}), |\eta| > 2.414 \quad (12)$$

$$histo(1) += Amp(\overrightarrow{D_{i,j}}), |\eta| \leq 0.414 \quad (13)$$

$$histo(3) += Amp(\overrightarrow{D_{i,j}}), 0.414 < |\eta| \leq 2.414 \quad (14)$$

따라서 위의 edge 방향을 구하는 식 (10)는 사용하지 않아도 되며 모든 화소 값에 대해 로버트 연산을 적용하여 edge의 방향에 따라 결정된 예측모드에 해당하는 edge의 크기 값을 누적하여 가장 큰 값을 갖는 예측 모드를 DC 모드와 함께 후보 모드로 선택한다. 이와 같이 결정된 후보 모드는 비트율 왜곡 최적 기법을 적용하여 두 모드 중 작은 값을 갖는 모드를 최적의 모드로 결정한다. 인트라 16×16 모드의 경우 4 가지 예측 모드 중 2 가지 모드에 대해서 비트율 왜곡 비용 값을 계산하므로 기존의 방법에 비해 모드 수를 현저하게 줄일 수 있는 효과를 얻는다.

반면에, 인트라 4×4 모드의 경우 인트라 16×16 모드와는 다른 방법으로 최적의 모드를 선택한다. 인트라 4×4 모드에서는 위에서 제시한 로버트 연산을 사용하지 않는다. 만약 로버트 연산을 사용할 경우 소벨 연산에 비하여 단순한 계산을 통해 화소간의 차이를 알아보는 것이기 때문에 4×4 블록 안에서 적용된다면 성능 저하를 초래할 수 있다. 따라서 인트라 4×4 모드에서는 Wang이 제시한 방향성과 같이 예측 모드 9 가지 중 4 모드, 즉 0모드(90°), 1모드(0°), 3모드(45°), 4모드(135°)를 대표 방향으로 설정한다. 전체 4×4 블록에서 2×2 블록씩 평행이동 하면서 그림 6에서 설정한 네 가지 방향에 대한 가중치를 이용하여 식 (15)와 같이 계산을 수행한다.

$$mode(S) = \left| \sum_{n=0}^3 a_n \times f^n \right|, \quad (15)$$

$S = 0^\circ, 45^\circ, 90^\circ, 135^\circ$ 의 윌터계수

4×4 블록에서 네 가지 모드에 대해 계산된 결과 중 0 모드와 1 모드 중에서 최대값을 갖는다면 선택된 모드를 예측 후보 모드로 정하고 선택된 예측 후보 모드와 가장 근접한 예측 모드 두 가지, 그리고 DC 모드도 함께 예측 후보 모드로 결정한다. 이렇게 선택된 네 가지 모드에 대해서만 비트율 왜곡 비

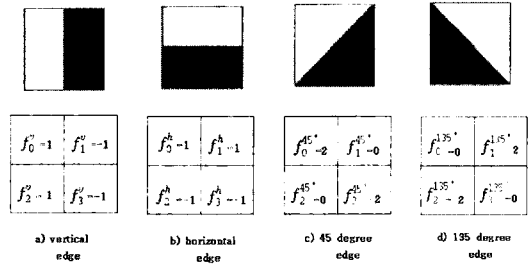


그림 6. 네 가지 방향성과 그에 따른 가중치 계수

표 3. 인트라 4×4모드에 예측후보모드 결정

인트라 모드	방향성 결정	예측 후보모드
인트라 4×4모드	0모드	0,5,7 과 DC
	1모드	1,6,8 과 DC
	3모드	0,1,3,7,8 과 DC
	4모드	0,1,4,5,6 과 DC

용 계산을 수행한다. 또한 3 모드와 4 모드 중에서 최대값을 가질 경우에는 인트라 4×4모드의 아홉 가지 예측 모드 중 선택된 모드와 같은 방향성을 가지는 모드 다섯 가지와 DC 모드에 대하여 비트율 왜곡 비용 계산을 수행하도록 한다. 따라서 다음의 표 3과 같이 인트라 4×4 모드의 경우 9 가지 예측 모드 중 4 가지 모드를 선택하거나 최대 6 가지 모드를 선택할 수도 있다.

IV. 모의 실험

본 논문에서 제안하는 알고리즘의 성능을 분석하기 위해서 H.264/AVC 참조 소프트웨어(JM 11.0)를 사용하여 실험하였으며 실험에는 Intel(R) Core(TM)2 2.13GHz CPU와 메모리 2GB를 갖는 컴퓨터를 사용하였다. 양자화 파라미터를 28로 하였고 CAVLC를 사용하였다. 제안하는 알고리즘의 성능평가를 위하여 네 개의 QCIF(4:2:0) 크기의 테스트 영상(Miss America, Carphone, Foreman, Claire)과 네 개의 CIF 크기의 테스트 영상(Stefan, Foot ball, Foreman, News)을 이용하여 실험을 하였다. 각 테스트 영상의 최초 10 프레임은 인트라 프레임으로 부호화 하였으며, Pan 등이 제안한 방법^[3]과 Wang 등이 제안한 방법^[4], 그리고 참조 소프트웨어(JM11.0)에서의 방법을 함께 비교하여 성능을 비교하였다.

표 4는 본 논문의 II장에서 분석했던 효율적인 인트라 예측모드 결정을 위한 기존의 방법에 대하여 원 영상과 비교한 PSNR 값을 나타내고 있다. 양자

표 4. 양자화 파라미터=28 일때 RDO사용 여부에 따른 PSNR 비교

size	sequence	RDO = ON				RDO = OFF			
		JVT	Pan's	Wang's	Proposed	JVT	Pan's	Wang's	Proposed
QC IF	miss_am	40.98	40.90	40.96	40.88	40.88	40.77	40.96	40.85
	carphone	38.21	38.15	38.16	38.21	38.05	37.97	38.03	38.06
	foreman	36.84	36.69	36.77	36.80	36.67	36.52	36.64	36.67
	claire	40.84	40.61	40.78	40.63	40.71	40.49	40.64	40.58
CIF	stefan	37.00	37.00	37.01	36.99	36.68	36.66	36.68	36.68
	football	37.30	37.21	37.24	37.27	37.15	37.06	37.11	37.15
	foreman	37.79	37.65	37.77	37.79	36.72	37.56	37.69	37.73
	news	39.01	38.94	38.97	38.96	38.82	38.78	38.80	38.82

표 5. JM11.0 소프트웨어에서의 비트스트림 비트율을 1로 하였을때 상대적 비트율 비교

size	sequence	RDO = ON				RDO = OFF			
		JVT	Pan's	Wang's	Proposed	JVT	Pan's	Wang's	Proposed
QC IF	miss_am	1.000	1.083	1.044	1.027	1.000	1.075	1.048	1.025
	carphone	1.000	1.061	1.037	1.015	1.000	1.057	1.047	1.008
	foreman	1.000	1.168	1.020	1.010	1.000	1.164	1.048	1.009
	claire	1.000	1.112	1.039	1.023	1.000	1.108	1.042	1.035
CIF	stefan	1.000	1.039	1.025	1.009	1.000	1.037	1.026	1.007
	football	1.000	1.072	1.033	1.012	1.000	1.070	1.042	1.008
	foreman	1.000	1.171	1.031	1.022	1.000	1.172	1.061	1.021
	news	1.000	1.075	1.051	1.021	1.000	1.071	1.052	1.018

화 파라미터 값으로 28을 사용했을 때 조건으로 비트율 왜곡 최적기법을 사용한 경우와 사용하지 않는 경우에 대해서 부호화 하였을 때 참조 소프트웨어 방법, Pan의 방법, Wang의 방법, 그리고 본 논문에서 제안하고 있는 방법에 대한 PSNR 결과를 나타낸다. 표 4와 표 5에서 제시한 PSNR과 비트율 결과를 보면 참조 소프트웨어에서 사용하는 인트라 16x16 모드의 4 가지 예측 모드와 인트라 4x4 모드의 9 가지 예측 모드에 대한 각각의 비트 왜곡을 계산할 수 행하는 것은 많은 계산량을 필요로 한다. 그에 반해 효율적인 알고리즘들이 제시한 방법은 예측 모드 수를 줄임으로써 수행시간을 크게 줄일 수 있었다. 그러나 속도가 향상된 만큼 예측 모드 수를 한정시켜서 인트라 예측 모드를 선택하는 것은 참조 소프트웨어에서 제시한 전체 예측 모드 중에서 선택하는 방법과 비교하였을 때 성능이 다소 저하된다. 본 논문에서 제안하는 방법은 인트라 16x16 모드와 인트라 4x4 모드 중 어떤 모드를 사용할지를 먼저 선택하고 더 효율적인 예측 모드 결정을 위해 로버트 연

산을 사용하였으며 4가지의 방향성 mask를 사용하여 예측 후보모드의 수를 감소시켰기 때문에 II장에서 분석했던 Pan등이 제안한 방법에 비해 계산량과 수행시간을 감소시켰다.

또한 기존의 방법에서 Pan 등이 제안한 방법은 예측 모드 수를 줄이기 위한 복잡한 소벨 연산 수행 과정 때문에 매크로 블록의 경계의 화소값들은 edge 성분을 정확하게 구하기 어렵다. 따라서 Pan 등이 제안하는 방법은 Wang 등이 제안한 방법과 본 논문이 제안하는 방법에 비하여 비트율이 높은 편이다. Wang 등이 제안한 방법은 인트라 예측 후보 모드를 9가지 모드 전체를 선택하여 비트율 왜곡 비용계산을 하는 경우도 존재한다. 이러한 점을 고려하여 본 논문은 방향성을 선택할 때 9 가지 모드를 후보 모드로 선택하는 경우는 사용하지 않았으며 방향성을 세밀하게 찾기 위해 4x4 블록에 존재하는 전체 화소값들을 이용하여 2x2 블록으로 평행이동을 하여 mask 연산을 수행하였다. 그 결과 비트율을 줄일 수 있었다.

마지막으로 표 6에 제시한 기존 알고리즘들은

표 6. JM11.0 소프트웨어에서의 전체부호화시간을 1로 하였을 때 RDO 사용 여부에 따른 상대적 시간(sec)비교

size	sequence	RDO - ON				RDO - OFF			
		JVT	Pan's	Wang's	Proposed	JVT	Pan's	Wang's	Proposed
QCIF	miss_am	1.000	0.841	0.794	0.447	1.000	1.155	0.923	0.808
	carphone	1.000	0.775	0.737	0.548	1.000	1.199	0.994	0.838
	foreman	1.000	0.793	0.786	0.576	1.000	1.152	1.000	0.850
	claire	1.000	0.839	0.754	0.439	1.000	1.20	1.038	0.956
CIF	stefan	1.000	0.768	0.765	0.541	1.000	1.061	0.966	0.850
	football	1.000	0.779	0.756	0.572	1.000	1.111	0.981	0.944
	foreman	1.000	0.824	0.777	0.532	1.000	1.120	0.962	0.860
	news	1.000	0.796	0.780	0.511	1.000	1.082	0.963	0.845

RDO의 사용여부에 관계없이 참조소프트웨어의 방법보다 수행시간이 감소하였으며 본 논문에서 제안한 알고리즘은 기존에 제시한 알고리즘에 비하여 수행시간이 더 크게 감소함을 알 수 있다.

V. 결 론

본 논문에서는 H.264/AVC 표준 부호화 기법에서 인트라 모드 결정 및 예측 모드 결정을 위한 효율적인 방법을 제안하였다. 인트라 모드를 결정하기 위해 분산 값을 사용하였고, 인트라 예측 모드 결정을 위해서는 인트라 16x16 모드의 경우 로버츠 연산을 이용하여 각 블록의 방향성을 계산하여 4 가지 예측 모드 중 2 가지 예측 후보 모드를 선택하고 인트라 4x4 모드의 경우 9 가지 예측 모드 중 4 가지 예측 후보 모드 또는 6 가지 예측 모드를 선택하였다. 화질 열화 및 비트율 증가는 무시할 수 있을 정도로 작은 반면에 RDO를 사용하였을 경우 기존의 효율적인 알고리즘에 비해서 평균 30% 이상의 속도향상을 보이며 RDO를 사용하지 않았을 경우 평균 10% 이상의 속도 향상을 보인다. 또한 기존에 제시된 방법에 비하여 비트율 또한 많이 감소시킴을 확인할 수 있었다. 이상에서 제안하는 방법은 H.264 표준을 위한 고속부호화에 적합한 효율적인 방법으로 사용될 수 있을 것이라 판단된다.

참 고 문 헌

[1] *Information Technology—Coding of Audio-Visual Objects—Part 10: Advanced Video coding*, Final Draft International Standard. ISO/IEC FDIS 14496-10, Mar. 2005.

[2] "Report of the formal verification tests on AVC(ISO/IEC 14496-10 ITU-T Rec. H.264)," MPEG2003 /N6231,dEC.2003.

[3] T. Wiegand, G. j. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.13, No.7, pp.560-576, July 2003.

[4] F. Pan, X. Lin, S. rahardja, K. P. Lim, Z. and G. Li, "A directional field based fast intra mode decision algorithm for H.264 video coding," *IEEE International Conference on Multimedia and Expo*, Vol.2, pp.1147-1150 June 2004.

[5] Jun Sung Park and Hyo Jung Song, "Fast selective intra mode decision H.264/AVC," *IEEE Consumer Communications and Networking Conference 2006. 3rd*, Vol.2, pp.1068-1072 Jan. 2006.

[6] An-Chao Tsai, Anand Paul, Jai-Ching Wang and Jhing-Fa Wang, "Efficient Intra Prediction in H.264 Based on Intensity Gradient Approach." *IEEE International Symposium on Circuit and Systems*, pp.3952-3955 May 2007.

[7] F. Pan, X. Lin, S. Rahardja, K. P. Lim, Z. G. Li, D. Wu and S. Wu, "Fast mode decision algorithm for intraprediction in H.264/AVC video coding," *IEEE Transactions on Circuits and Systems for Video Technology*, Vol.15, No.7, pp.813-822, July 2005.

[8] J. F. Wang, J.C. Wang, J. T. Chen, A. C. Tsai and A. P, "A novel fast algorithm for

intra mode decision in H.264/AVC encoders.”
IEEE International Symposium on. Circuits and systems, pp.3498-3501, May 2006.

- [9] H. Kalva and B. Petljanski, “Exploiting the directional features in MPEG-2 for H.264 intra transcoding,” *IEEE Transactions on consumer Electronics*, Vol.52, No.2, May 2006.
- [10] J. B. Song, B. Li, W. Li and L. Jiang, “A Novel Fast Intra Prediction Algorithm applied in H.264/avc,” *IEEE International Conference on Signal Processing*, Vol.1, pp.16-20 2006.

박 지 윤 (Ji-yoon Park)

준회원



2008년 2월 가톨릭대학교 정보통신전자공학부 졸업
2008년 3월~현재 가톨릭대학교 정보통신전자공학과 석사과정
<관심분야> 영상처리, 신호처리, 영상통신

이 창 우 (Chang-woo Lee)

정회원

현재 가톨릭대학교 정보통신전자공학부 교수
<관심분야> 영상통신, 영상처리